

SEQUENCE LISTING

<110> Kakefuda, Genichi
Costello, Colleen
Sun, Ming
Hu, Weiming

<120> Genes and Vectors for Conferring Herbicide Resistance
in Plants

<130> 043753/241148 (5849-20A)

<140>

<141>

<150> 60/106,239

<151> 1998-10-29

<150> 09/426,568

<151> 1999-10-22

<160> 11

<170> PatentIn Ver. 2.1

<210> 1

<211> 1673

<212> DNA

<213> Arabidopsis sp.

<220>

<221> CDS

<222> (42)..(1514)

<223> Mature Peptide

<400> 1

gtcttcttca gtagcaaaaa accttcggct tcgtctcgta a atg gcg gcc att tct 56
Met Ala Ala Ile Ser
1 5

gta agt tct tca cca tct att cgc tgc ttg aga tcg gca tgt tcc gat 104
Val Ser Ser Ser Pro Ser Ile Arg Cys Leu Arg Ser Ala Cys Ser Asp
10 15 20

tct tct cct gct ctt gta tcc tcg acg cgt gta tcg ttc ccg gcg aag 152
Ser Ser Pro Ala Leu Val Ser Ser Thr Arg Val Ser Phe Pro Ala Lys
25 30 35

att tca tat ctc tcc ggt ata tct tcg cac cgt ggc gat gaa atg ggt			200
Ile Ser Tyr Leu Ser Gly Ile Ser Ser His Arg Gly Asp Glu Met Gly			
40	45	50	
aag aga atg gaa gga ttc gtt aga agc gtc gat ggg aag atc tct gat			248
Lys Arg Met Glu Gly Phe Val Arg Ser Val Asp Gly Lys Ile Ser Asp			
55	60	65	
gcg tct ttc tcc gaa gct tca tct tcg act cca aaa tcg aag gtg agg			296
Ala Ser Phe Ser Glu Ala Ser Ser Ala Thr Pro Lys Ser Lys Val Arg			
70	75	80	85
aag cac aca att tca gta ttt gtt gga gac gaa agc gga atg att aat			344
Lys His Thr Ile Ser Val Phe Val Gly Asp Glu Ser Gly Met Ile Asn			
90	95	100	
agg att gca gga gtg ttt gca agg aga gga tac aat att gag agt ctt			392
Arg Ile Ala Gly Val Phe Ala Arg Arg Gly Tyr Asn Ile Glu Ser Leu			
105	110	115	
gct gtt ggt ctg aac aga gac aag gct cta ttc acc ata gtt gtc tgt			440
Ala Val Gly Leu Asn Arg Asp Lys Ala Leu Phe Thr Ile Val Val Cys			
120	125	130	
gga act gaa agg gta ctt cag cag gtc atc gag caa ctc cag aag ctc			488
Gly Thr Glu Arg Val Leu Gln Gln Val Ile Glu Gln Leu Gln Lys Leu			
135	140	145	
gtt aat gtt cta aag gtt gaa gat atc tca agt gag ccg caa gtg gag			536
Val Asn Val Leu Lys Val Glu Asp Ile Ser Ser Glu Pro Gln Val Glu			
150	155	160	165
cgt gag ctg atg ctt gta aaa gtg aat gca cat cca gaa tcc agg gca			584
Arg Glu Leu Met Leu Val Lys Val Asn Ala His Pro Glu Ser Arg Ala			
170	175	180	
gag atc atg tgg cta gtt gac aca ttc aga gca aga gtt gta gat ata			632
Glu Ile Met Trp Leu Val Asp Thr Phe Arg Ala Arg Val Val Asp Ile			
185	190	195	
gcg gaa cat gca ttg act atc gag gta act gga gat cct gga aaa atg			680
Ala Glu His Ala Leu Thr Ile Glu Val Thr Gly Asp Pro Gly Lys Met			
200	205	210	
att gct gta gaa aga aat ttg aaa aag ttt cag atc aga gag att gta			728
Ile Ala Val Glu Arg Asn Leu Lys Lys Phe Gln Ile Arg Glu Ile Val			
215	220	225	

agg aca gga aag ata gca ctg aga agg gaa aag atg ggt gca act gct			776
Arg Thr Gly Lys Ile Ala Leu Arg Arg Glu Lys Met Gly Ala Thr Ala			
230	235	240	245
cca ttt tgg cga ttt tca gca gca tcc tat cca gat ctc aag gag caa			824
Pro Phe Trp Arg Phe Ser Ala Ala Ser Tyr Pro Asp Leu Lys Glu Gln			
250	255	260	
gcg cct gtt agt gtt ctt cga agt agc aaa aaa gga gcc att gtc cct			872
Ala Pro Val Ser Val Leu Arg Ser Ser Lys Lys Gly Ala Ile Val Pro			
265	270	275	
caa aag gaa aca tca gca ggg gga gat gtt tat ccc gtt gag cca ttt			920
Gln Lys Glu Thr Ser Ala Gly Gly Asp Val Tyr Pro Val Glu Pro Phe			
280	285	290	
ttt gac ccc aag gta cat cgt att ctc gac gct cac tgg gga ctt ctc			968
Phe Asp Pro Lys Val His Arg Ile Leu Asp Ala His Trp Gly Leu Leu			
295	300	305	
act gac gaa gat acg agt gga cta cg ^g tcg cat act cta tca ttg ctt			1016
Thr Asp Glu Asp Thr Ser Gly Leu Arg Ser His Thr Leu Ser Leu Leu			
310	315	320	325
gta aat gat att cca gga gtt ctt aat att gtg act ggt gtt ttc gct			1064
Val Asn Asp Ile Pro Gly Val Leu Asn Ile Val Thr Gly Val Phe Ala			
330	335	340	
cga agg gga tac aat atc cag agc ttg gcc gta gga cat gct gaa acc			1112
Arg Arg Gly Tyr Asn Ile Gln Ser Leu Ala Val Gly His Ala Glu Thr			
345	350	355	
aag ggc att tca cgc att aca aca gtt ata cct gca aca gat gaa tcg			1160
Lys Gly Ile Ser Arg Ile Thr Thr Val Ile Pro Ala Thr Asp Glu Ser			
360	365	370	
gtc agc aaa ttg gtg cag caa ctt tac aaa ctc gta gat gtg cat gag			1208
Val Ser Lys Leu Val Gln Gln Leu Tyr Lys Leu Val Asp Val His Glu			
375	380	385	
gtc cat gat ctt act cat ttg cca ttt tct gaa aga gaa ctg atg ctg			1256
Val His Asp Leu Thr His Leu Pro Phe Ser Glu Arg Glu Leu Met Leu			
390	395	400	405
att aag att gcc gtg aac gct gct gct aga aga gat gtc ctg gac att			1304
Ile Lys Ile Ala Val Asn Ala Ala Arg Arg Asp Val Leu Asp Ile			
410	415	420	

gct agt att ttc agg gct aaa gct gtt gac gta tct gat cac aca att			1352
Ala Ser Ile Phe Arg Ala Lys Ala Val Asp Val Ser Asp His Thr Ile			
425	430	435	
act ttg cag ctt act ggg gat cta gac aag atg gtt gca ctg caa agg			1400
Thr Leu Gln Leu Thr Gly Asp Leu Asp Lys Met Val Ala Leu Gln Arg			
440	445	450	
tta ttg gag ccc tat ggt ata tgt gag gtt gca aga acc ggt cgt gtg			1448
Leu Leu Glu Pro Tyr Gly Ile Cys Glu Val Ala Arg Thr Gly Arg Val			
455	460	465	
gca ttg gct cgt gaa tcg gga gtg gac tcc aag tac ctt cgt gga tac			1496
Ala Leu Ala Arg Glu Ser Gly Val Asp Ser Lys Tyr Leu Arg Gly Tyr			
470	475	480	485
tcc ttt ctt tta aca ggc taaaccgttg cagagtgcac ccatcgaaca			1544
Ser Phe Leu Leu Thr Gly			
490			
tcagaaactt tggaaggtaa aagtttcatt acacagtcta tgaacctcaa agacagacag			1604
agagactgcg tcgatatatg tttgtgactt tgtttatgaa acaattagct gatTTggc			1664
ttcatttcg			1673
<210> 2			
<211> 491			
<212> PRT			
<213> Arabidopsis sp.			
<400> 2			
Met Ala Ala Ile Ser Val Ser Ser Ser Pro Ser Ile Arg Cys Leu Arg			
1	5	10	15
Ser Ala Cys Ser Asp Ser Ser Pro Ala Leu Val Ser Ser Thr Arg Val			
20	25	30	
Ser Phe Pro Ala Lys Ile Ser Tyr Leu Ser Gly Ile Ser Ser His Arg			
35	40	45	
Gly Asp Glu Met Gly Lys Arg Met Glu Gly Phe Val Arg Ser Val Asp			
50	55	60	
Gly Lys Ile Ser Asp Ala Ser Phe Ser Glu Ala Ser Ser Ala Thr Pro			
65	70	75	80

Lys Ser Lys Val Arg Lys His Thr Ile Ser Val Phe Val Gly Asp Glu
85 90 95

Ser Gly Met Ile Asn Arg Ile Ala Gly Val Phe Ala Arg Arg Gly Tyr
100 105 110

Asn Ile Glu Ser Leu Ala Val Gly Leu Asn Arg Asp Lys Ala Leu Phe
115 120 125

Thr Ile Val Val Cys Gly Thr Glu Arg Val Leu Gln Gln Val Ile Glu
130 135 140

Gln Leu Gln Lys Leu Val Asn Val Leu Lys Val Glu Asp Ile Ser Ser
145 150 155 160

Glu Pro Gln Val Glu Arg Glu Leu Met Leu Val Lys Val Asn Ala His
165 170 175

Pro Glu Ser Arg Ala Glu Ile Met Trp Leu Val Asp Thr Phe Arg Ala
180 185 190

Arg Val Val Asp Ile Ala Glu His Ala Leu Thr Ile Glu Val Thr Gly
195 200 205

Asp Pro Gly Lys Met Ile Ala Val Glu Arg Asn Leu Lys Lys Phe Gln
210 215 220

Ile Arg Glu Ile Val Arg Thr Gly Lys Ile Ala Leu Arg Arg Glu Lys
225 230 235 240

Met Gly Ala Thr Ala Pro Phe Trp Arg Phe Ser Ala Ala Ser Tyr Pro
245 250 255

Asp Leu Lys Glu Gln Ala Pro Val Ser Val Leu Arg Ser Ser Lys Lys
260 265 270

Gly Ala Ile Val Pro Gln Lys Glu Thr Ser Ala Gly Gly Asp Val Tyr
275 280 285

Pro Val Glu Pro Phe Phe Asp Pro Lys Val His Arg Ile Leu Asp Ala
290 295 300

His Trp Gly Leu Leu Thr Asp Glu Asp Thr Ser Gly Leu Arg Ser His
305 310 315 320

Thr Leu Ser Leu Leu Val Asn Asp Ile Pro Gly Val Leu Asn Ile Val
325 330 335

Thr Gly Val Phe Ala Arg Arg Gly Tyr Asn Ile Gln Ser Leu Ala Val
340 345 350

Gly His Ala Glu Thr Lys Gly Ile Ser Arg Ile Thr Thr Val Ile Pro
355 360 365

Ala Thr Asp Glu Ser Val Ser Lys Leu Val Gln Gln Leu Tyr Lys Leu
370 375 380

Val Asp Val His Glu Val His Asp Leu Thr His Leu Pro Phe Ser Glu
385 390 395 400

Arg Glu Leu Met Leu Ile Lys Ile Ala Val Asn Ala Ala Arg Arg
405 410 415

Asp Val Leu Asp Ile Ala Ser Ile Phe Arg Ala Lys Ala Val Asp Val
420 425 430

Ser Asp His Thr Ile Thr Leu Gln Leu Thr Gly Asp Leu Asp Lys Met
435 440 445

Val Ala Leu Gln Arg Leu Leu Glu Pro Tyr Gly Ile Cys Glu Val Ala
450 455 460

Arg Thr Gly Arg Val Ala Leu Ala Arg Glu Ser Gly Val Asp Ser Lys
465 470 475 480

Tyr Leu Arg Gly Tyr Ser Phe Leu Leu Thr Gly
485 490

<210> 3
<211> 4895
<212> DNA
<213> *Arabidopsis* sp.

<220>
<221> promoter
<222> (1)..(757)
<223> Promoter Region

<220>
<221> misc_feature
<222> (717)
<223> Transcriptional Starting Point

<220>

<221> misc_signal
<222> (758)..(760)
<223> Start Codon

<220>
<221> misc_signal
<222> (4737)..(4739)
<223> Stop Codon

<220>
<223> n at position 694 can be a, c, g, or t

<400> 3

tcgcataattg ttccggcgag gatcatgtga agcttgacgc gtgaattgac gactaagcgt 60
acgacgaagc gatccagttg agaattgtct cgagattcct cgttttagct gtcccactac 120
attcgccatg atttcgaaat ctcttctct tcttctct ttcgtcttct tctgcgaaaa 180
aatcgaatgg ataatcacat ttcttttc tcgagaaaaat tgatctggtg attatgtgag 240
atccgtctct agcgcgttgc ttatcgagaa ataattaatt ttaatttgac gggtaagat 300
attattggcg acgtctgtt ccgattgact ttgatttgac tttcccttcc aatcattatt 360
tggcgagtc ccgcgtaaata tggactcttc ttgattgtcc cacttttc ggtggcttta 420
ccggatttaa aatcattttc tttccttaaa ttatgaattt taccctaaac ttctcataat 480
tacaatttagt tccgacgaac ccaagatact ttttagcaaa attagaaaa tagttgactc 540
gaaaaggttgg tataacgtg gagctgacgt gttggctta tctactcgaa gccttttggg 600
cttttcttaa agccattgtat ttctaaggtc gtcaacaacc gaaccggacc ggacggtttgc 660
accggtctaa ccaacatata tacgttctt ttcnacttgc cgttcgtcg tcgtcagtct 720
tcttcagtag caaaaaacct tcggcttcgt ctctgtcaatg gcggccattt ctgtaaagtcc 780
ttcaccatct attcgctgct tgagatcggtc atgttccgat tcttctccgt ctcttgcattc 840
ctcgacgcgt gtatcggtcc cggcgaagat ttcatatctc tccggatataat cttcgacccg 900
tggcgatgaa atgggtaaga gaatggaaagg attcggttgcgatg ggaagatctc 960
tgatcgctt ttctccgaag cttcatctgc gactccaaaa tcgaagcgac tgtgaataat 1020
atttgcttaa agtcgtttcc ttttggcatt tgctttgatt gattcttgcgt gcattaaat 1080
cagggtgagg aagcacacaa tttcagttt tggtggagac gaaagcggaa tgattaatag 1140
gattgcagga gtgtttgcaaa ggagaggata caatattgag agtcttgcgt ttggctcgaa 1200
cagagacaag gctctattca ccatagttgt ctgtggact gaaagggtac ttcaagcagg 1260
catcgagcaa ctccagaagc tcgttaatgt tctaaagggtt gttcttgcgt tagatcgac 1320
ttatttagttt ctgcattgtact atagtttcat tcgcaccaac tttacgcattt agccaatttgc 1380
cttattcatt atttgaagat tagatttgcg atttcctttt ccattcttctt cattgacttgc 1440
gacatgaatt aggttgaaga tatctcaagt gagccgcaag tggagcgtga gctgtatgtt 1500
gtaaaaagtga atgcacatcc agaatccagg gcagaggtac tattccttgc ctatggaaaa 1560
tttaggtta ctgtacttgc tgggtgcattc tgatattttt cagaggtgtt gtttagtttgc 1620
tctctaaatt tgatattttt tctgtttttaa tgaattcaca gatcatgtgg ctatgtaca 1680
cattcagagc aagagttgtt gatatacggtt aacatgcattt gactatcgat gtcacatctac 1740
ttattatgtt ttgtgttgcgtt ctgtatattt gtttgcactt gtagccttgc ggtttcaaga 1800
cttctgtttt aacatcttac taatcggttgg aagacatcgaa aatattatgtt gagggtatcat 1860
tttaactttt atatcttattt gttggatattt cgttgcattt tgaaactgtat gatgtatccac 1920
atgcaggact ctattatagg atgtgttattt aagtttattt gaaacttttg gtgcacttgc 1980
ttgaatttaa tataacgaga aagtttattca acagtgtgtt acctttgattt accctatgtt 2040
tataatctgtt attctgatgtt gtttgcattt tgcaaaatttc ttttttttttttgc 2100

acttttggaaa gttagagaag cataacccta aatatattgt tcttttacc ttgattatga 2160
gaaagtggag taaaagaaaag ggtgtctctg atttacctat tttagctctt tagtaatcat 2220
tttaagcta ttttcaggt aactggagat cctggaaaaa tgattgctgt agaaagaaaat 2280
ttgaaaaagt ttcagatcg agagattgt aaggacaggaa aggtatgtt tgtttggaaat 2340
tactagattt tatggctttt gaatatcatc tagttgtgc tatctaattgt atgtatgttag 2400
tagttacatc tttgagtgga cacaaggca tagatctcag ggacttcac taatttaggg 2460
aaaatggaaat gacattttg gataacagat agcactgaga agggaaaaga tgggtgcaac 2520
tgctccattt tggcgatttt cagcagcatc ctatccagat ctcaggagc aagcgcctgt 2580
tagtggctt cgaagtagca aaaaaggagc cattgtccct caaaaggaaa catcagcagg 2640
ggtgtgtgt tctctgctcc ttagattgtt taacttcagc ttgaagttcc tcactttcc 2700
ttcaaaaaat ttggttgcat aaattatagt aggtttggct atttgataaa gttaaacagc 2760
aactatagat gcctgtgtt tttccctct atgtggtggc tgcctggaat caacatttga 2820
agcatgcct tttttgttt tctccctggc tgactgaag gattccgag tttgctaatt 2880
tttaaaagtt atcttatctt tttaaatgt aggagatgtt tatcccgtt agccattttt 2940
tgaccccaag gtacatcgta ttctcgacgc tcactggga cttctcactg acgaagatgt 3000
aagagagttc ttgctatata atctaacttc gtgccatgaa tttgctaaaa agcaatatga 3060
aaaattcaga ttgtggttt cattacacga gttacacttgc ttttccattt caagccgtct 3120
ggcataatca attctgttaa tatagttaca taaatgataa atcaatttgc tgtagattt 3180
ggagactgta tgtatattact tacaaggcag acattgaaag agttggggtt ttctttaagc 3240
tatttcgttt tatttatcac agttattttt ttttgcattt tcagacgagt ggactacgg 3300
cgccatactt atcattgcattt gtaaatgata ttccaggagt tcttaatattt gtgactgggt 3360
ttttcgctcg aaggggatac aatatccagg catagtcctt atctctctca tacacgcaca 3420
cacagtgtgc ttaggtttac tgacacactg aaagatctcc tttcttttag agcttggccg 3480
taggacatgc tgaacccaag ggcatttcac gcattacaac agttatacc gcaacagatg 3540
aatcggtcag caaatgggt cagcaactt acaaactcg agatgtgc tggatgggat 3600
tacccaaagc tactgtctttt cttatataattt taacagtttgc aatgtctttt atggccctat 3660
cattcccttgc ctgtcttaga cctttggct tttttaaaaa cgttagatttgc aggaagagtt 3720
tctgctaaat ctttctggac tttcctatattt catttcctgg tcttgcgtt tttactcgat 3780
gagaccttgc ttccagaaa gtcacactgt acaggcttgc tgaaaataat tctgaacatg 3840
atttgcgcac actttccaag ctgttattaa ctttgcggg attttctgc ggtccatgt 3900
cttactcatt tgccattttc tgaagagaaa ctgatgcgtt ttaagattgc cgtgaacgct 3960
gctgctagaa gagatgtctt ggacatttgc agtatttca gggctaaagc ttttgacgt 4020
tctgatcaca caattactt gcaggtaaaa tacatttctc ataaatgggat tttttatgt 4080
gctgttatttgc catctcagat gagaatccctt ttcaatttgc gatcttcaaa gtttgcacgtc 4140
tttccatagg tcttcaactt gtttgcata atcagatgttgc cttttggaaa aaatataatga 4200
agctgacttgc gattttccat cttaatcttctt tttttttgc ttttgcgtt ggattttgtt 4260
gctgaaattt gttggctgtt ggtatagctt actggggatc tagacaagat ggttgcactg 4320
caaaggat tggagcccta tggatataatgtt gaggtttgc tgcacatcta ctttcatctc 4380
tttagtgaatg cataaccccg tgaatttcta ttttttataa tgctacccca attgctccgg 4440
ataaaagtccc aaaatttatgt tttttttatgtt acgacttgc aacagatgt tgaacatcta 4500
actctctgtt aaaatcaata accaaagctg gacctgttgc catgaatctt cttctgggtt 4560
tgtgtagaac aagaataagc ttgacaagcc atgactactt tcagattatgc catcggttt 4620
acgcttatttgc tgaacatca atcacacagg ttgcaagaac cggtcgtgtt gcattggctc 4680
gtgaatcggtt gttggactcc aagtacccttgc tggatactc ctttcttttgc acaggctaaa 4740
ccgttgcaga gtcacatccat cgaacatcgaa aactttggat ggttggaaaat ttcattacac 4800
agtctatgttgc cctccaaagac agacagagat gtcgttgc tttatgttttgc tgactttgtt 4860
tatgaaacaa ttagtgcattt ttgggttgc tttcg 4895

<210> 4
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer made
from cDNA sequence

<220>
<221> misc_feature
<222> (1)..(22)
<223> DNA primer

<400> 4
cagagatcat gtggctagtt ga

22

<210> 5
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer made
from cDNA sequence

<220>
<221> misc_feature
<222> (1)..(22)
<223> DNA primer

<400> 5
gagcgtcgag aatacgatgt ac

22

<210> 6
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Thrombin
cleavage site

<400> 6
Leu Val Pro Arg Gly Ser

<210> 7
<211> 6
<212> PRT
<213> Arabidopsis sp.

<220>
<221> PEPTIDE
<222> (1)..(6)
<223> N-terminal of AHAS small subunit peptide of
pHUWE82

<220>
<221> SITE
<222> (1)..(2)
<223> Thrombin cleavage site

<400> 7
Gly Ser Ile Ser Val Ser
1 5

<210> 8
<211> 6
<212> PRT
<213> Arabidopsis sp.

<220>
<221> PEPTIDE
<222> (1)..(6)
<223> N-terminal of AHAS small subunit peptide of
pHUWE83

<220>
<221> SITE
<222> (1)..(2)
<223> Thrombin cleavage site

<400> 8
Gly Ser Met Ile Asn Arg
1 5

<210> 9
<211> 8
<212> PRT

<213> Arabidopsis sp.

<220>

<221> PEPTIDE

<222> (1)..(8)

<223> N-terminal sequence of AHAS small subunit peptide
of plasmid F1

<220>

<221> SITE

<222> (1)..(2)

<223> Thrombin cleavage site

<400> 9

Gly Ser Pro Lys Ile Ala Leu Arg

1

5

\

<210> 10

<211> 7

<212> PRT

<213> Arabidopsis sp.

<220>

<221> PEPTIDE

<222> (1)..(7)

<223> N-terminal sequence of AHAS small subunit peptide
of plasmid F2

<220>

<221> SITE

<222> (1)..(2)

<223> Thrombin cleavage site

<400> 10

Gly Ser Leu Asp Ala His Trp

1

5

<210> 11

<211> 7

<212> PRT

<213> Arabidopsis sp.

<220>

<221> PEPTIDE

<222> (1)..(7)

<223> N-terminal of AHAS small subunit peptide from

plasmid F3

<220>
<221> SITE
<222> (1)..(2)
<223> Thrombin cleavage site

<400> 11
Gly Ser Val Glu Pro Phe Phe
1 5